

Student Assessment in Adult Basic Education: A Canadian Snapshot

Written by Dr. Pat Campbell
February 2006

ACKNOWLEDGEMENTS

The purpose of this document is to report on the findings from a national survey on student assessment in adult basic education. This report is one of several outcomes from a national project entitled *Assessment Practices in Adult Basic Education*. This project will also produce an edited book on assessment practices and three videos.

This project was sponsored by the Centre for Education and Work (CEW) in Winnipeg, Manitoba. The CEW gratefully acknowledges the financial support provided by the National Literacy Secretariat, Department of Human Resources and Social Development Canada for funding this project.

We would like to acknowledge the contributions of the advisory committee, who helped to form the survey framework and questions:

Suzanne Benoit
Tom Ciancone
Anne Marie Downie
Guy Ewing
Karen Geraci
Lynda Fownes
Laurie Gould
Diana Kaan
Robin Millar
Beverly Sabourin
Cate Sills

A special thank-you to the 400 educators who took the time to participate in the survey. Your comments and insights made the survey come alive!

Finally, thank you to Yvan Clermont, Statistics Canada; Todd Rogers, Centre for Research in Applied Measurement and Evaluation at the University of Alberta; and Grace Malicky for providing suggestions on improving the on-line survey.

For more information on this project, contact

Dr. Pat Campbell
Project Director
PO Box 52192
Edmonton, AB
T6G 2T5

Tel: 780-448-7323
Fax: 780-413-6582
pcampbell@interbaun.com

Table of Contents

Introduction.....	5
Methodology.....	7
Assessment Tools	11
Purposes Of Assessment.....	29
Assessment Practices	37
Professional Development	49
Muddy Water.....	55
Key Findings and Recommendations.....	63
References	69
Appendices.....	75



INTRODUCTION

Assessment is a continuous, systematic process that utilizes tools and approaches to gather information in order to make decisions about the provision of programs, instruction, training, and services for literacy, upgrading and adult basic education students. Adult basic educators who work in programs offered by colleges, school boards, workplaces, and community-based agencies use a variety of assessments that are distinguished by the theories on which they are based; the knowledge, skills, attitudes, and values they purport to measure and provide; and the manner in which they inform placement and instruction. Although assessment is an integral part of the learning process, there is a paucity of information about the assessment tools and practices that are being used by adult educators across Canada. In addition, the research on the supports and constraints that affect educators' ability to engage in effective assessment practices is sparse.

Since the turn of the 21st century, the face of literacy provision and assessment has changed, in response to changing views about the nature of literacy and numeracy learning. In order to ensure that existing assessment tools reflect our changing views and meet the needs of educators and the diverse range of students, a national survey on student assessment was conducted with 400 educators in 2005. This report unfolds with a brief description of the methodology used to conduct the national survey. This report documents the findings from the survey, with the intent that the findings be used to inform policy and practice.



METHODOLOGY

Development of Survey

In November, 2004, an 11-member national advisory committee met to design the survey framework and questions. A draft survey was developed and forwarded to the advisory committee and statisticians at Statistics Canada and the Centre for Research and Applied Measurement and Evaluation (CRAME) at the University of Alberta for feedback. The survey was revised, based on this feedback. An on-line version of the survey was programmed and piloted with eight adult literacy educators. The findings from the pilot study were used to revise the survey, and the final version was completed in March 2005 (see Appendix A).

Data Collection

Governments and coalitions from the 13 jurisdictions provided contact information for English-language literacy and adult basic education programs offered by three kinds of delivery agencies: college, school board, and community. If a delivery agency in a given jurisdiction had less than 30 programs, all of them were asked to participate in the survey. For example, in the territories and the Atlantic Provinces, every program was contacted, since fewer than 30 programs were offered by each type of delivery agency. In order to ensure a representative sample, 50 percent of the programs were randomly sampled whenever a delivery agency in a given jurisdiction had more than 30 programs. For example, in Ontario, only 50 percent of the programs offered by colleges, school boards, and community-based programs were contacted, as each type of delivery agency had more than 30 programs.

The survey was conducted between March and August 2005. The goal was to achieve a 75 percent response rate. The director's initial contact with the potential survey respondent was made by phone. In many instances, the initial contact person redirected the director to another educator in the program who held more responsibility for assessment. During the phone call, the purpose of the survey was discussed and the

writing, and numeracy; and to a lesser extent, instruction in life, employment, and essential skills. The respondents indicated that the four primary modes of instruction are one-to-one (75 percent), small group (65 percent), self-paced (63 percent), and class (54 percent). Given the high percentages for each mode, it is apparent that many programs use a combination of these modes.

ASSESSMENT TOOLS

Assessment Tools Used by Delivery Agencies

A major purpose of the survey was to inventory the English-language assessment tools being used by delivery agencies and to identify the strengths and limitations of these tools. The findings indicate that the 400 respondents use 26 different types of commercial instruments to assess literacy, numeracy, and essential skills. Appendix B lists the tools that each jurisdiction uses for initial assessment.

Of these 26 instruments, only three are current instruments that were developed within the past decade for the adult Canadian population and were being used at the national level. These three tests, Canadian Adult Reading Assessment (*CARA*), *Common Assessment of Basic Skills (CABS)*, and Test of Workplace Essential Skills (*TOWES*), were funded by the National Literacy Secretariat. A total of 134 respondents use *CARA* for initial, on-going, and/or exit assessment, making it the most widely used test of the 26 different instruments. *CABS* was the second most widely used test, with a total of 81 respondents using it for the various stages of assessment. *TOWES* was used only by 53 respondents, but this is due to the fact that it was developed *specifically* for the Canadian workplace to assess essential skills in three domains. Since only 51 of the 400 respondents identified themselves as workplace educators, it is not surprising that *TOWES* was used only by 53 respondents. Clearly, *TOWES* is being used for the purpose for which it was intended.

Despite the range of instruments being used across the country, the findings portray an assessment profile for the different delivery agencies and preferences for specific tools *across* the delivery agencies. In adult basic education, assessment tools and tasks that measure and document performance can be classified into four types: authentic, diagnostic, standardized, and competency-based. The findings portray a specific assessment profile for each delivery agency; they favour certain types of assessment tools over others. For example, credit-granting institutions such as community colleges, school boards, and Manitoba's learning centres favour standardized tools over diagnostic

Diagnostic Assessments

Community-based and workplace programs, unlike credit granting institutions, tend to favour diagnostic assessments over standardized and competency-based assessments. Diagnostic instruments provide instructors with the student's reading levels; this information leads to appropriate initial placement in classes and/or materials. Diagnostic tests require strong interpretive skills that can be gained only through a combination of experience and training. The interpretation of the test results provides specific information about how individuals process print and text, and identify specific areas of weakness and strength in a learner's skills and capabilities. This information is used to design effective lesson plans, form instructional objectives, and determine reading strategies that will meet these objectives.

The Canadian Adult Reading Assessment.

The findings show that *CARA* is the most frequently used diagnostic assessment across delivery agencies.¹ *CARA* consists of a 186-page instructor's manual, a 52-page student's assessment booklet, and a CD-ROM. The instructor's manual and student's assessment booklet contain a graded word list and nine levels of passages, ranging in readability from Grades 1 to 12. *CARA* provides instructors with the student's independent, instructional, and frustration reading levels.

The *CARA* levels are correlated with the adult literacy and adult basic education programs offered through each province and territory. Each level contains two fiction and three non-fiction passages. *CARA* has a high degree of alternate-form reliability. This means passages within each level are parallel in terms of difficulty. Consequently, *CARA* can be used to measure progress and evaluate gains in reading level, processes, and strategies.

A total of 121 (30 percent) respondents reported that *CARA* is one of the tools used for initial assessment. Of the 121 respondents, 53 stated that *CARA* was the *most*

¹ It should be noted the the author of this paper is also the co-developer of *CARA*. This should be taken into account when reading the *CARA* review as the author might be biased towards this tool.

frequently used tool for initial assessment. The 53 respondents described the strengths and limitations of *CARA*.

In terms of strengths, the respondents cited the dual purposes of *CARA*: it provides information on both placement *and* instruction. The respondents stated that *CARA* provides reliable and accurate placement and “usually is right on the button with the placing.” The respondents described *CARA* as an “easy-to-use, flexible” tool that provides students with choices in terms of passage selection. The passages were cited as a strength for several reasons: (1) Canadian content; (2) high-interest and culturally appropriate stories; and (3) the range in difficulty, from beginning to advanced. Several respondents appreciate that *CARA* is articulated to the levels recognized in the different provincial and territorial jurisdictions.

The findings show that *CARA* allows the assessor “to connect with the student” and establish rapport. This is because *CARA*’s administration process allows for “interaction, communication, and feedback” between the assessor and student. Consequently, the students find the assessment “non-threatening” and are not “intimidated.”

The most frequently cited limitation was that *CARA* is time-consuming, particularly with students who have well-developed reading skills. Second, *CARA* was criticized because it is “not able to test readers at extremely low levels.” Third, the graded word list was critiqued for several reasons: (1) Level 1 list is too difficult; (2) two of the words—“hymn” and “yacht”—are culturally biased; and (3) there is no parallel form for the word lists. Although the title—*Canadian Adult Reading Assessment*—implies a focus on reading, many respondents expressed a desire for *CARA* to assess the domains of numeracy and writing.

Authentic Assessments

Authentic assessment may be conducted with an array of assessment texts and tasks that are usually informal measures developed by educators. The majority of educators across all the delivery agencies stated that they used tools developed by practitioners for initial (68 percent), on-going (72 percent), and exit assessment (51 percent). The assessment tasks require the application of knowledge or skills in contexts that reflect literacy and numeracy practices within the home, work, community, and school. Typically, authentic assessment occurs during regular classroom or tutoring activities, which means that educators can provide immediate and responsive feedback to students. Sometimes, authentic assessment materials are adopted by programs within other regions or jurisdictions.

Since authentic assessments are developed by educators, each one tends to be unique. Due to the limitations of space, only three authentic assessments used for initial assessment in different types of delivery agencies are described. The following description portrays an authentic assessment used for initial assessment of literacy skills. This comprehensive assessment is used by a college instructor, with a masters degree and more than 20 years experience in the field of adult basic education:

Our literacy assessment has four components: interview, phonics, reading, and writing. We begin with the interview, talking with the student about what brought them back to school, what goals they have, what their interests are, etc., and explaining what our program offers. The phonics assessment is very simple, just asking them to identify single consonants, consonant blends, and give the sound and a word beginning with that sound. Students are also asked to read a few consonant-vowel-consonant words, and a few silent “e” words. If the phonics assessment indicates that the student is a complete non-reader, we use a book of simple signs and ask them to tell us about what the signs are saying. Students usually find the assessment encouraging because almost everyone can identify all the letters, so everyone experiences some success with it.

For the reading portion, students are usually given a selection of several articles (e.g., from the *West Coast Reader*) and asked to choose one they think they can read at least some part of. After the reading we go up or down with the next article, depending on how the first reading went. We

are looking for something that will give us a sense of their reading level and the types of strategies they use to decode. When the students have read the articles, we also ask comprehension questions and/or ask students to tell us about the article, allowing them to refer back to it.

For the writing portion, students are asked to write/print the alphabet, complete sentences with a word(s) left blank, answer a few questions in writing (e.g., What food do you like to eat?), and then write a few sentences (or more, if possible) about a place they have lived, or any other topic they would like to write about.

We stop the assessment at any point when it is clear that the student will not be able to handle the next level or task.

This assessment reflects literacy practices within the classroom. Several colleges have developed an assessment that “can be used to determine a student's instructional reading level, writing strengths and weaknesses, and knowledge of mathematical concepts and operations.” For example, the Basic Education Department at Vancouver Community College has published an intake assessment tool that is being used by several colleges in British Columbia. In the territories, Aurora College has developed the Aurora College Placement Test, and Nunavut Arctic College has designed placement tests for Math, Inuktitut, and English.

Workplaces tend to develop contextualized performance assessments that assess the essential skills used in the workplace or occupation. A field officer who coordinates workplace education programs provided the following description of the initial assessment for workers in a specific occupation:

The process is informal in that it is carried out in an interview with the learner and a range of activities are offered and the learner can self-assess and then choose the activity that he/she feels comfortable with. For example: for uncertified workers in the field of early childhood we used a growth chart and asked the participants to determine the weight/height of a two-year-old if he were in the 50th percentile for his age. We refer to the Essential Skills profiles for ideas. The tasks have a complexity level assigned and it is easy to duplicate that activity with materials from the workplace to create the needs assessment.

The essential skills framework was used by workplace educators who work with many different populations. For example, one respondent works “in a sheltered workshop for clients with special challenges.” One of the tools she developed uses essential skills related to the retail store. The assessment results are used to “customize the training program to the right skill level of the participants.”

A practitioner with a master’s degree and over ten years of experience in adult literacy provided the description of an initial assessment. This practitioner works in a community-based program in the inner city.

When I interview potential students at the Centre, I use the Centre's registration form as an assessment tool. The form definitely has meaning for students and is used in context. It provides me with enough information to place the student in a group or with a tutor—more specific needs can be addressed later when the student is more at ease. I usually invite the student into my office, where I invite them to sit and have a look at the registration form while I continue with a task in my office. I am busy with a task so the student does not feel they are being watched (although I am aware of their actions); almost all of the students will complete as much of the form as they can while they wait. After an appropriate length of time, I turn my full attention to them, and we go over the form together. Blanks, spelling, questions asked—these allow an entry assessment (not exact to be sure) without the stress of using an 'assessment tool.' It's not rocket science, but it works!

This description encapsulates why many practitioners choose to use authentic assessment over a commercial tool. Students speak with their feet; if the initial assessment creates stress and anxiety, they might not return to the program. This practitioner, like many others, is concerned with all aspects of a student’s well-being; this includes his/her emotional state and not just the cognitive results from an assessment.

A total of 267 (67 percent) respondents reported that authentic assessments are part of the initial assessment battery. Of the 267 respondents, 125 stated that tools developed by instructors were the most frequently used tools for initial assessment. The 125 respondents described the strengths and limitations of these tools.

According to the findings, authentic assessment has two key strengths. First, these tools are customized to meet the needs of specific programs and students. The tasks and activities can reflect learning outcomes, curriculum, and/or course content, which means that instructors can use the test results for placement into classes or to develop individualized programs. The content can reflect the essential skills used in workplace environments, which means that instructors can use the test results to develop training programs for the individual learner. The tools can also be tailored to a student's cultural background, goals, and interests. Finally, the tools can be developed for special needs students, such as those with intellectual disabilities.

Second, the respondents emphasized that authentic assessments do not intimidate students who are returning to school. The tools are flexible instruments that allow the assessor and student to choose test items that are relevant to the student. Moreover, they are simple and easy to administer and interpret. The high level of flexibility, choice, and interaction between the assessor and the student creates a relaxed setting for the student. The informal nature of the assessment means that the students feel comfortable and are more "willing to take risks." Reducing the student's anxiety and alleviating the stress related to tests "increases the likelihood that the assessment accurately represents the students skills."

The respondents cited five main limitations with respect to the use of authentic assessment. First, these assessment tools provide general information about the student's capabilities, but do not provide a detailed assessment of the learner's abilities that would enable educators to "narrow down the specific areas the student needs to work on." Further to this, one respondent noted that authentic assessment "doesn't capture the actual source of learning problems. Rather, it tends to point at symptoms rather than underlying causes." For example, the assessment might reveal that the student has difficulties with comprehension, but might not unmask the source of these difficulties. Moreover, these assessments provide only a rough indication of performance level, making it difficult to use test results for placement purposes. Several respondents stated these assessments are limited in scope; the tools tend to assess reading and writing

competencies, but do not fully assess other areas such as numeracy and speaking and listening skills. In summary, the general nature of authentic assessment tools impacts upon their ability to be used for placement and instructional purposes. As well, these tools cannot be used to measure progress in terms of reading levels, as they do not contain parallel forms.

Second, the respondents expressed the concern that these tools were not recognized by other institutions. This means that the assessment results have limited transferability between institutions. Consequently, these tests cannot always be used for exit assessments.

Third, effective administration and interpretation of authentic assessment is dependent upon the experience of the examiner. According to one respondent, the “the potential weakness of [authentic assessment] is that it requires an experienced educator to evaluate the work produced by the student and to begin his/her program at the appropriate level.” Some educators, particularly those who are new to the field, may not possess the observational skills needed to identify a student’s strengths and weaknesses. It is often difficult for educators to access generic training in how to use these tools, as each tool is unique to the agency that developed it.

Fourth, the interpretation of test results can be subjective. This can result in differences of opinion about how to place and instruct the student. One respondent stated that “the weakness is in the interpretation of the results among instructors, especially with writing assessments. There is sometimes a difference of opinion about what level the writing sample is at.”

Finally, many of the respondents commented that authentic assessments are time-consuming, in terms of development and administration. It is particularly challenging for part-time instructors to find the time to develop these tools. In summary, educators need to weigh the time it takes to develop and administer these tools against the information the tools provide in terms of placement and instruction.

Standardized Assessments

Credit-granting institutions such as colleges, school boards, and Manitoba’s learning centres tend to favour standardized assessments over diagnostic and competency-based assessments (see Figure 3). The term “standardized” means that the tests are always expected to be administered and scored the same way. The majority of standardized tests are “norm-referenced.” This means that the test compares a student’s scores to the scores of a representative or norm group.

The Canadian Adult Achievement Test.

The *CAAT* is ranked as the most frequently used standardized assessment across delivery agencies. The *CAAT* measures an adult’s current functional level in math, reading, and language. It consists of a battery of nine sub-tests: reading vocabulary, reading comprehension, spelling, number operations, problem solving, mechanical reasoning, language, science, and study skills. There are four levels of *CAAT*, with each level corresponding to years of education and featuring a different combination of subtests. Level A, for instance, is for adults who have completed 1 to 3 years of formal education, while Level D is for adults who have completed 11 to 12 years of formal education. *CAAT*, which was normed on adults, provides norm-referenced and content-referenced scores.²

A total of 83 (22 percent) respondents reported that *CAAT* is one of the tools used for initial assessment. Of the 83 respondents, 54 stated that *CAAT* was the *most frequently* used tool for initial assessment. The findings clearly indicate that the 54 respondents do not always agree about *CAAT*’s strengths and weaknesses.

Many respondents appreciated the fact that *CAAT* provides grade equivalency and stanine scores. The grade equivalency scores were used to determine placement in classes, while the stanine scores were used to determine areas of strengths and weaknesses across a range of domains. Many cited reliability as one of *CAAT*’s major strengths, stating that

² It should be noted that the standard definition of a reliable test is one that yields consistent test scores over time and different test situations, so many respondents were introducing a slightly modified definition of reliable.

“CAAT has proven reliable in the placement of adult learners” into appropriate adult basic education (ABE) classes.¹ This finding was contradicted by others, who found that “the results do *not* give an accurate indication of the reading/numeracy ability of the learner.” Specifically, these respondents found that the grade level scores on CAAT were too high, which can result in “inappropriate grade placement.” Some respondents surmised that the multiple-choice format of the test might be a factor in inflated test scores, as the student could “guess” the correct answer.

Many respondents found that CAAT is easy to administer and score and doesn’t require the staff to have specialized credentials. Moreover, the respondents find it to be a cost-effective instrument because it can be administered in a group setting. However, others complained that the administration takes too long and the “test-taker” becomes fatigued.

The findings revealed that the respondents did agree on four limitations. First, CAAT can have a negative impact on the student’s emotional state. The respondents used the terms “intimidating,” “scary,” and “overwhelming,” to describe CAAT and went as far to say it can “create anxiety for the learner.” Second, CAAT appears to have a “southern urban bias” and is “not culturally sensitive.” Third, CAAT is not the “best tool for the assessment of lower level students.” Fourth, many respondents reported that the writing sub-test is a major weakness because it does not require the individual to provide a writing sample. In fact, a writing subtest is offered only in Levels C and D; students are provided with four sentences that express the same idea and are asked to choose the one which is written clearly and correctly.

Competency-based Assessments

Approximately one in five respondents who worked in credit-granting institutions and one in four respondents who worked in community-based programs and workplaces use competency-based assessments (see Figure 3). Competency-based assessments are designed to measure skills and knowledge as they relate to predetermined competencies, which can include life, workforce, and/or academic skills. This type of assessment is intended to measure the individual’s specific literacy, numeracy, and/or essential skills in order to determine what he/she knows and can do in *specified contexts*. The individual’s

performance is mapped against a set of predefined, specified skills within a competencies matrix. Criterion-referenced tests can also be classified as competency-based assessments because the student's performance is measured against a standard or curricular content. Criterion-referenced assessment is linked to instructional objectives, and individual items are designed to assess mastery of specific objectives.

It should be noted that authentic assessment and competency-based assessment are similar in terms of the nature of some of the tasks. The significant difference between the two types of assessment is that the competency-based assessment evaluates the student's performance against a set of predefined skills or outcomes. A further difference is that authentic assessments are typically developed by instructors, whereas most of the competency-based assessment tools, such as the *Tests of Workplace Essential Skills (TOWES)* and *Common Assessment of Basic Skills (CABS)*, are commercial products.

Common Assessment of Basic Skills (CABS).

CABS is ranked as the most frequently used competency-based assessment. *CABS* is designed to align the individual's skills against Ontario's five level *Learning Outcomes Matrix*.³ This information can be used for placement into appropriate classes and for the choice of appropriate materials. *CABS* can be administered only to an individual, never to a group.

The 380-page *CABS* manual outlines a four-step process for the initial contact, interview, assessment, and follow-up of students. The manual provides an interview guide, forms, and sample demonstrations. The interview guide and assessment are also available online.

CABS uses demonstrations to show competencies in communications (reading, writing, speaking, and listening) and numeracy at the five ability levels established in the matrix. Demonstrations are activities linked to learning outcomes that resemble real-life tasks that are related to the learner's goals. Each demonstration is accompanied by an assessor's

³ The five levels are being articulated to the levels established in other provinces and territories.

guideline page. After each completed demonstration, the assessor compares the student's work to performance indicators and success markers listed on the page. This descriptive scale helps the assessor to determine the student's level(s) and competencies. The *CABS* manual contains 94 demonstrations, with 17 to 19 demonstrations at each of the 5 levels. There have been no studies to establish psychometric measures such as reliability and validity.

A total of 71 (18 percent) respondents reported that *CABS* is one of the tools used for initial assessment. Of the 71 respondents, 26 stated that *CABS* was the *most frequently* used tool for initial assessment. The 26 respondents described the strengths and limitations of *CABS*. It should be noted that 25 of these 26 respondents were from Ontario.

The respondents like *CABS* because it serves three purposes. First, *CABS* establishes a student's communications and numeracy levels, based on Ontario's five-level *Learning Outcomes Matrix*. This information can be used for placement into appropriate classes and for the choice of appropriate materials. Second, "the strength of *CABS* is its relevance to the reporting we are required to do to our Ministry." *CABS* meets the accountability demands of the Ontario Ministry of Training, Colleges, and Universities because it aligns with Ontario's *Learning Outcomes Matrix* and has been approved by the Ministry of Training, Colleges, and Universities (MTCU). Third, the test results are portable when the student transfers to a different program because *CABS* is a provincially recognized tool used by delivery agencies across Ontario.

In terms of its features, respondents stated that *CABS* is a user-friendly tool that is easy to administer. The respondents appreciate the breadth of demonstrations in the manual, as this allows them to choose activities tailored to the student's goals and interests. The respondents like the demonstrations because they mirror real-life activities. Finally, some of the respondents commented that *CABS* is a non-threatening instrument because it is "not too formal, not too much like a test, and uses real life examples."

While many respondents commented on the breadth of demonstrations, others stated that they would “like to see more” in the next edition. A frequent request was for demonstrations that are “workplace-related.” The respondents recognized that *CABS* is time-consuming and “often takes two sessions to complete.” Yet, they felt that the time “was well-spent because it provides a way to get to know the learner’s background and goals for the future.”

Tools for the Future

What type of assessment tools would you like to see developed in the future? This question resulted in an extensive wish list from 330 respondents, ranging from tools that assessed workforce and academic readiness to those that diagnosed math anxiety. Despite this range, data analysis indicated that the highest rankings were for assessment tools that assess (1) emergent literacy and numeracy skills; (2) numeracy and writing skills; and (3) learning disabilities. The respondents wanted the tools to be user-friendly and culturally sensitive for First Nations and English as a Second Language (ESL) students.

Emergent Literacy and Numeracy Assessment Tools

Among the respondents, the highest-ranked need was for a tool that could be used with students whose literacy and numeracy skills and processes are at a very basic level. The respondents wanted a user-friendly tool that would not intimidate the learner or the assessor, as many programs utilize volunteers to assess students. They wanted this tool to measure progress, which means that it must include alternate forms or components for initial, on-going, and exit assessment. They also wanted this tool to inform instruction, which means that reference material that shows how the assessment results link to instruction would be beneficial. Several respondents remarked that they wanted to use this tool with ESL students, as well as literacy students.

Existing assessment tools do not provide a comprehensive assessment of students with limited literacy and numeracy skills, making a tool of this nature a necessity. Currently, three commercial Canadian assessment tools designed for adults are being used across the country: *CAAT*, *CARA*, and *CABS*. Each of these tools possesses particular strengths and

weaknesses in terms of its ability to assess the literacy and numeracy of emerging readers and writers. For example, *CARA* does not provide tools to assess an emergent reader's repertoire of sight words and phonemic awareness or his/her phonics knowledge. Further, it is not a comprehensive tool in that it assesses only reading.

Numeracy and Writing Assessment Tools

There are several standardized, diagnostic, and competency-based tools that contain numeracy and writing subtests. This raises the question, “Why did the respondents perceive the need for yet another instrument to assess these areas?” The primary reason is that the respondents want a user-friendly, culturally sensitive, and non-threatening tool that focuses on the assessment of *one* skill—numeracy or writing. In other words, they want an instrument that provides an in-depth assessment, rather than a subtest that deals with writing or numeracy in a limited manner.

Numeracy.

The respondents identified the need for a tool that measured numeracy levels, skills, processes, and conceptual understanding. They called for a tool that used *authentic* tasks, with practical, everyday applications. The respondents wanted a tool that could measure all levels of math, from basic math to college level.

One respondent noted, “I would like a numeracy assessment tool that was more than a page of problems—most tools are too difficult for the learners, and do not provide enough information to plan effective instruction.” While some wanted the tool to address learning challenges and inform instruction, others wanted the tool to assess academic readiness and determine placement in programs. One respondent expressed the need for a quick screening tool to determine placement. Some respondents wanted a standardized tool normed on adults, while others wanted a diagnostic tool that could pinpoint strengths and weaknesses.

Writing.

The respondents wanted a creative, interesting writing assessment that actually required the student to write. One instructor noted that “many of our assessments ask students to label the types of errors made in sample sentences or ask the student to choose a correct answer from multiple choices.” These tasks cannot be considered valid measures, as they do not involve the competency being assessed: writing.

The respondents called for a tool that could assess the student’s key strengths and weaknesses, and consequently inform instruction. For example, one respondent wanted a tool that referenced the strategies she needed to teach in order to address the student’s weaknesses. Although the administration of a writing assessment can be straightforward, the scoring and interpretation is not. The respondents asked for guidelines that would help them to interpret writing samples and provide feedback to the student.

Learning Disabilities

When a student has difficulty learning, an immediate response is “Why?” Quite often, educators believe that learning challenges stem from an undetected learning disability. So it was no surprise to learn that many of the survey respondents wanted an assessment tool capable of diagnosing learning disabilities.

In order to be diagnosed as learning disabled, individuals need to meet four criteria. The individual must:

- Have average to above-average intelligence.
- Demonstrate a discrepancy between intelligence and achievement test scores.
- Demonstrate a discrepancy between achievement and age and ability.
- Have learning problems that are not primarily the result of visual, hearing, or motor disabilities; mental retardation; emotional disturbance; or environmental, cultural, or economic disadvantage.

A battery of psycho-educational tests must be administered to diagnose these criteria. One of these criteria, having average to above-average intelligence, can be determined only through an intelligence test such as the *Stanford-Binet Intelligence Scale* or the *Wechsler*

Adult Intelligence Scale – Revised. In order to administer these scales, one needs to be a certified psychologist or have the appropriate credentials.

Many respondents wanted an “assessment tool to identify possible learning disabilities and appropriate instruction.” The follow-up report from a psycho-educational assessment battery does not always provide the specific reading strategies that a student needs to learn in order to become a fluent reader. This may be because the assessor is usually a psychologist, rather than a reading specialist. A diagnostic assessment tool, on the hand, is usually an effective means of identifying the cognitive processes a student uses to decode words and comprehend text. This type of assessment can be administered by a reading specialist or an educator. While diagnostic tools do not provide a label, such as “learning disabled,” they do provide the information, such as specific strategies, needed to effectively instruct a student. The advantage of a psycho-educational assessment, on the other hand, is that it usually provides learners with the “type of documentation needed to access accommodations in post-secondary environments.”

In summary, although the survey respondents wanted an assessment tool capable of diagnosing learning disabilities, it is simply not feasible to invest in its development. This type of diagnosis requires a battery of tests administered by people with the appropriate credentials. There exists, however, a screening tool for adults at risk of possible learning disabilities. It was designed by the Learning Disabilities Association of Canada (LDAC) in the early 1990s and further refined in 1999 in the publication *Destination Literacy: Identifying and teaching adults with learning disabilities*. This screening tool can be used to predict “the possibility of a learning disability...Although lacking the status of a proper assessment, it is highly practical in identifying an individual’s learning style where accommodations in work/educational settings might be appropriate.”

Over the past 20 years, I have facilitated dozens of workshops on assessment across the country. Educators in these workshops continually express frustration about being unable to diagnose students who are at risk of being learning disabled. Perhaps, as educators, we need to shift our gaze and reframe our perspectives on this issue. Rather than focusing on our inability to diagnose learning disabilities, we could shift our gaze to the strategies a

student uses for processing print and text. Using a psycho-educational battery to diagnose learning disabilities requires programs to access certified professionals and funding, whereas using a diagnostic assessment to figure out how a student processes print and text simply requires the programs to provide educators with the training they need to administer and interpret diagnostic assessments. Once adult educators have the training and confidence, they can use diagnostic tools to determine a student's learning strengths and challenges, and use this information to develop appropriate instruction.

PURPOSES OF ASSESSMENT

A frequently asked question among educators is “What is the best tool to use in our program?” The answer to this question depends partially on the purpose of the assessment. The survey findings indicate that, from the perspective of educators, the primary purposes of assessment are its potential to determine placement (80 percent), inform instruction and training (70 percent), and/or establish a benchmark or skill level (62 percent). Only 21 percent of the respondents recorded that the purpose of assessment is to measure progress.

The assessment tools and approaches should be appropriate for and compatible with the purpose and context of the assessment. For example, if the purpose is to determine placement, the educator should use a valid and reliable instrument aligned with current theories and practices. If the purpose is to inform instruction, the tool should provide in-depth information about how the student processes print and text. If the purpose is to establish a benchmark or skill level, the tool should be aligned to provincial, territorial, or national standards. Finally, if the purpose of assessment is to measure progress, the tool should have parallel forms and reflect classroom practice. Unfortunately, educators cannot always choose tools appropriate for the given purpose because educational policies and constraints often dictate the tools that must be used.

The survey asked two questions to explore the purposes of assessment. The first question listed the purposes of initial assessment and asked the respondents whether their program rated each purpose as primary, secondary, or tertiary. The second question asked the respondents to select the three most important factors in choosing an assessment tool. This section discusses whether the assessment tools that respondents use are appropriate for and compatible with the purpose and context of the assessment.

Determining Placement

The findings clearly indicate that for the majority of respondents (80 percent), the primary purpose of initial assessment in their program is to determine the student’s placement.

PURPOSES OF ASSESSMENT

Moreover, the highest-ranked factor for choosing an assessment tool is the tool's ability to provide information for placement; 57 percent of the respondents selected this factor. Given that the respondents view determining placement as one of the primary purposes of assessment, educators and policy-makers need to ensure that the assessments they use or mandate accurately determine the student's level of proficiency.

In an ideal learning environment, assessment tools used for determining placement would meet the following criteria. They would be valid and reliable assessment tools that reflect current literacy and numeracy theories and curriculum. Moreover, they would be normed on an adult population and free of bias. Many assessments do not accurately portray the student's levels of proficiency because they fall short of this set of criteria. One respondent stated that the assessment tool used in Nunavut is "proving to be inadequate in assessment placement or insuring that the student is being placed properly. This is affecting student success in our program." Another respondent from British Columbia commented that "it is very stressful for students and instructors when someone is placed at a higher grade level than indicated by the assessment." Even though educators are aware of the limitations of the tools they use, they continue to use them for several reasons. For example, 50 respondents (13 percent) reported that they are mandated to use specific assessment tools. Out of the 50 respondents, 25 stated that the mandated tools they use are inappropriate. Other educators reported they do not have the time or funds to research and learn about the range of tools and approaches that are available.

Validity

The document *Principles for Fair Student Assessment Practices for Education in Canada* (1993), states that validity is enhanced when the assessment tool reflects instructional goals and objectives and is compatible with instructional approaches. Yet, many tests used for determining placement do not reflect the curriculum goals and objectives. For instance, the *Canadian Achievement Survey Test (CAST)* was criticized by one of the respondents because the "math portion is not reflective of the new math curriculum. The math test places students at a higher level than they should be placed, based on the new pure math curriculum." According to a review of this test, there is no validity data for

CAST, and “users must rely on their judgment of the face validity of the items” (Daniel, 2001). If the skills that a test measures do not align with what is being taught and learned in the adult basic educational system, caution should be exercised if the test is being used to determine placement.

Reliability

Reliability refers to the degree to which test scores are consistent over time and different test situations. The length of a test is one of the factors that influences reliability; tests with larger numbers of test items produce higher reliabilities. One of the respondents who used the *Brigance Diagnostic Reading Inventory of Essential Skills* stated that “as an initial placement tool I don't feel that it gave a very accurate assessment. The sampling of questions was too small.” Indeed, in the past, the *Brigance* has been criticized by a review in the *Mental Measurements Yearbook* for its lack of psychometric data on reliability (Matuszek, 1985). The reliability of CABS is also questionable because the directions do not specify how many demonstrations should be completed in each content domain. This can result in inadequate sampling. In a worst-case scenario, an educator might engage the learner in only a few demonstrations to determine placement. One of the respondents stated that some of the CABS demonstrations may not “depict the actual LBS (Literacy and Basic Skills) level.”

Norming Populations

Fifty-one respondents (13 percent) who administer assessments to determine placement are using instruments that have been standardized and normed on youth, rather than on an adult population. For example, the *Canadian Achievement Test (CAT)*, the *Canadian Test of Basic Skills (CTBS)*, and the *Gates-MacGinitie Reading Tests* were normed on students in the K-to-12 public school system. In a nutshell, tests normed on children and adolescents are not an accurate measure of adults' literacy and numeracy skills. Despite their obvious limitations, these tests continue to be used for placement. In fact, only a small percentage of the respondents (15 percent) who use tools for determining placement ranked the norming population as an important factor in test selection.

PURPOSES OF ASSESSMENT

A total of 39 respondents (10 percent) use *CAT* for initial assessment. Two respondents who use *CAT* remarked that “it can give false assessment of skills—usually lower than normal” and that “the written section does not necessary accurately reflect the student's placement.” According to a recent review (Soares, 2005), *CAT-3* is a well-constructed measure of basic academic skills in reading and mathematics. However, this test was not designed for the adult population; it was designed for students aged 11 to 20 years old. The norms, which were developed from a sample of 44,529 students from the public school system, are expressed as a grade equivalent score. Therefore, it is highly likely that *CAT-3* will give a distorted assessment of skills for the adult population.

The *Canadian Test of Basic Skills—Forms 7 and 8 (CTBS)* was used by only three respondents (1 percent). An educator who works in a college stated that the *CTBS* “does not always accurately reflect the student’s skills or abilities (i.e., sometimes weaker students score higher than their present skill level).” Another respondent who worked in a college that uses the *CTBS* commented, *CTBS* “has been used for many years—we understand that it doesn't test adults well and use it simply for initial placement.” The fact that only three educators use this tool does not sound very significant. But, when one considers the number of students who have been assessed over the years with this tool, one realizes the impact of a single tool used by a few people.

Some institutions continue to use the same tool, despite their awareness of the tool’s limitations. Although there are no data to verify this assumption, one could assume that some educators develop an emotional attachment to the tests they use. The educators are cognizant of the limitations of a test, yet they continue to use the test because they are familiar with its administration, scoring, and interpretation. This familiarity, in turn, creates a level of comfort with the test.

Bias

Bias occurs in testing when items systematically measure differently for ethnic, gender, or age groups. Many of the respondents commented that the tests they used contained cultural bias, particularly towards First Nations and ESL students. One respondent

noted that “the *CAT II* has cultural biases that do not measure First Nations' traditional knowledge, and generally First Nation students place at a lower level than necessary with the *CAT II*.” If educators use assessments that contain bias towards specific populations, the students' scores could well be deflated and not reflect their true abilities.

Kline (2004) recognized that “as with most tests of cognitive ability, questions are raised regarding potential for bias” (p. 550). Kline analyzed the relationship between subscale test performance on the *Test of Workplace Essential Skills (TOWES)* and the demographic variables of gender and language. Kline found “differences on all of the mean *subscale scores* for language groups. In each case, the non-English group scored about 10 percent lower than did those who had English as their first language” (p. 558). It should be noted that the *items* were not biased in favor of one gender or language when Kline used a differential item functioning (DIF) analysis. However, Kline concludes by stating that the “statistically significant difference in the estimated ability levels suggests that whether English is the first language of the test-taker should be a consideration for item development and interpretation of subscale performance” (p. 558). What does this mean for educators? The implication of Kline's study is that if educators think a test contains bias, they should probably adjust the test scores accordingly if the results are going to be used to determine placement.

Informing Instruction and Training

A student's assessment protocol can be used to develop responsive programs, design lesson plans, choose materials, and implement effective teaching strategies. The degree to which assessment can inform instruction depends on the qualifications and experience of the assessor and his/her ability to interpret assessments. One respondent noted: “Many students are not receiving proper instruction because we are not trained to assess their real learning problems. It is often a hit or miss.” Adult educators need “proper training on assessment tools and support to develop learning materials based on the scoring and interpretation of assessment results.” The initial assessment provides educators with a snapshot of the student's capabilities that is tied to a time and place, while ongoing assessment provides an unfolding montage of reading patterns and behaviours. On-

going assessment also keeps students abreast of their progress and determines whether they are benefiting from instruction; if the assessment indicates little or no progress, then instruction needs to be adjusted.

The assessment battery can be a combination of formal measures, such as commercial assessment tools, and informal approaches such as interviews, checklists, anecdotal records, observation, conferencing, demonstration activities, miscue analysis, portfolio assessment, learning self-assessment guides, peer assessment, and writing samples. The findings indicate that the respondents use a battery of formal and informal measures, with higher percentage of the respondents using informal approaches to inform instruction, in comparison to formal measures. A total of 92 percent of the respondents use commercial tools for initial assessment and 72 percent for on-going assessment. In comparison, 99 percent of the respondents use informal approaches for initial assessment and 98 percent for on-going.

Informal Approaches

Informal approaches have the potential to unveil the complicated realities of learning. Observation, for example, can be a very powerful assessment tool. Let's say you notice that a student prefers to "stick to the facts" rather than "read between the lines." Informal approaches allow educators to interact with the student to explore this phenomenon at a deeper level. Commercial assessments, on the other hand, would have a straightforward answer to this phenomenon: the student needs instruction in making inferences. However, through informal approaches such as conferencing, you might discover that the student does not realize he/she can exert power over the text and make inferences. Conversations with students about "reading the lines" and "reading between the lines" can help educators to understand the degree of power that students assert over the text.

Some of the informal approaches were used by the majority of the respondents, while others were not. For example, 80 percent of the respondents engaged in observation for on-going assessment, while peer assessment was used by only 16 percent of the respondents (see Figure 4).

PURPOSES OF ASSESSMENT

that the primary *purpose* of initial assessment is to identify the student's strengths and weaknesses. Yet only 39 percent of the respondents reported that a tool's ability to assess strengths and weaknesses was an important *factor* in choosing an assessment tool. This set of findings leads one to hypothesize that educators who want assessment tools to inform instruction might be choosing inappropriate tools.

The data confirmed this hypothesis. A total of 294 respondents indicated that in their program, the primary purpose of initial assessment was to plan for instruction and/or identify a student's strengths or weaknesses. However, only 102 of these respondents actually choose and use assessment tools that are aptly suited for this purpose. In other words, only one in three educators are using tools capable of meeting their stated purpose: to inform instruction. In most cases, the respondents are using tools that are better suited for determining placement. For example, 22 of the 294 respondents (7 percent) relied solely on one commercial instrument—*CAAT*—for initial assessment. This standardized test is outdated and does not reflect current views of the reading process; consequently, it does not provide enough information to adequately inform the teaching and learning process.

The International Reading Association is also concerned about tests that define reading as a set of discrete skills, rather than a set of practices and processes, especially when these tests are used to inform instruction. The Delegates Assembly of the International Reading Association adopted the following resolution:

Reading assessment must reflect advances in the understanding of the reading process. As teachers of literacy we are concerned that instructional decisions are too often made from assessments which define reading as a sequence of discrete skills that students must master to become readers. Such assessments foster inappropriate instruction. (p. 1)

As educators, we need to ensure that the assessment tools we use match the purpose for which they were intended.

ASSESSMENT PRACTICES

Effective Practices and Procedures

The respondents were provided with a list of best practices and asked to select the ones that their program followed. The two practices that received the highest ranking were directly related to protecting the student. Eighty-seven percent of the respondents reported that their assessments are conducted in a non-threatening manner, which is a way of *protecting* the student’s emotional well-being. The importance of non-threatening assessments was a recurring theme, woven into the responses to open-ended questions throughout the survey. Another respondent wrote:

I think the most critical issue is that we need to value the students and not place them in assessment situations where they are at risk of suffering humiliation...I think that nationally, we need to address this issue, because I am aware that some institutions simply test the students using the *CAT* test and the results are either acceptance into a program or rejection.

For many respondents, the testing environment is one of the most critical issues that need to be addressed at the national level.

Eighty-four percent of respondents indicated that the student is ensured of confidentiality, which is a way of *protecting* his/her privacy. While the respondents are cognizant of the importance of confidentiality, only 47 percent indicated that their program has a written policy to ensure confidentiality when the assessment protocol is transferred to another program. According to the *Principles for Fair Student Assessment Practices for Education in Canada* (1993), “transfer of assessment information from one school to another should be guided by a written policy with stringent provisions to ensure the maintenance of confidentiality” (p. 13).

The practice of sharing assessment results with the student received the third highest ranking (82 percent). The respondents provided details about the type of information that is shared with the student after the initial assessment (see Figure 5). In addition to sharing

“What is the point of engaging in professional development on assessment if one does not have the time to utilize what he/she has learned?” Professional development is effective only when practitioners have the time to practice, dialogue, and reflect upon their new knowledge. Simply put, until the issue of capacity is addressed, professional development on assessment will not lead to more effective practice. The words of one respondent sum up the dilemma: “The Ministry expects us to do it (assess), but never provides enough funding.” Funders need to ensure that educators have the capacity to respond to what is learned through professional development.

One woman commented on the “huge time factor involved in planning appropriate ongoing and exit assessments.” This, coupled with the fact that many students leave mid-way through the program, without providing notice, makes it difficult to use assessments to monitor progress. It is also challenging to make assessment a priority when there are so many competing responsibilities, duties, and pressures that consume and impinge upon an educator’s time. The following statement, from the director of an adult basic education program in a small rural college, represents the multi-faceted roles of many practitioners who work in rural, urban, *and* remote locations:

I feel that my initial assessments are good but since I am responsible for every aspect of the program from administration, assessment, training, tutor training, matching, goal setting, plans, information and referral, etc., I find that my ongoing and exit assessments are therefore sometimes lacking.

The issue of access surfaced as the second highest ranked support *and* constraint. A high percentage of the respondents have access to expertise or resource people (43 percent); access to journals, articles, and reports on assessment (42 percent); and access to literature accompanying assessment tools (30 percent). On the other hand, obstacles to effective assessment practice included isolated working conditions (27 percent) and access to resources (20 percent). Living in a geographically isolated community means that access to colleagues, professional development, support, and assessment tools is limited. Practitioners who live in urban centres also experience isolation, simply because

networking opportunities have not been established or are limited. One respondent from Winnipeg noted:

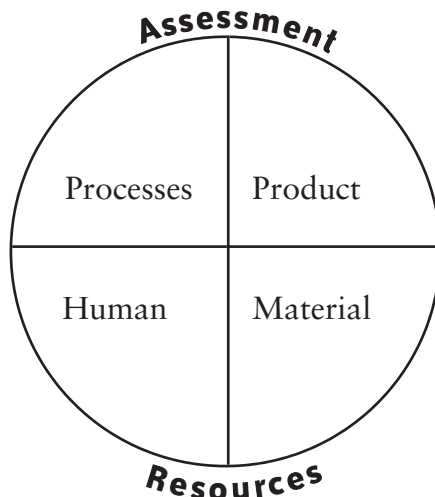
We seem to work so much in isolation. We [program co-ordinators] do meet, but this has only just started this year and we have not addressed the issue of assessment. Basically, we all just seem to do our own “thing” and make our own decisions based on our own knowledge and skill set.

While assessment can be learned through trial and error, assessment is also a socially constructed phenomenon that needs to be learned through dialogue and reflection with colleagues. It should be noted that limited access to resources is not restricted to practitioners living in isolated areas; in fact, many educators in urban centres noted the need for a library of assessment tools.

Advocating for Change

If you could wave a magic wand, what changes, if any, would you make to assessment practices in your program? This question generated interest, resulting in feedback from 335 respondents. Data analysis revealed that change revolved around two key areas: assessment processes and products, and human and material resources (see Figure 6). Although these two areas are presented as separate entities, they are connected, forming a gestalt. For example, additional human and material resources are usually a prerequisite to changes in assessment processes and improvements in assessment products.

Figure 6.



The Assessment Process

Comprehensive assessments.

Many respondents expressed a desire to administer comprehensive, in-depth intake assessments with individuals, rather than conducting group assessments. They wanted to analyze and interpret the assessment protocols in order to make informed decisions about instruction and design learning plans tailored to the individual's needs. Further, they wanted time to discuss the assessment results with the students, and provide an opportunity for students to ask questions. The data indicated that time was the primary barrier preventing people from conducting comprehensive assessments and providing feedback to the student.

Stages of assessment.

Assessment can strike fear into the hearts of students because tests symbolize their negative experience in the K-to-12 school system. Yet intake assessment continues to be the first step in the registration process for many upgrading programs. Several survey respondents did not want the initial contact to include assessment because it can discourage prospective learners, and it “puts up barriers and resistance.” One woman, who worked in a community-based program, wrote “I would allow a longer ‘get to know you’ time frame before the assessment testing is completed.” Another respondent who worked in the correctional system wanted “a process where the inmate would be stabilized before being assessed.” The practical considerations of postponing assessment, however, are particularly difficult in colleges dealing with large intakes of students; in these situations, determining placement in an adult basic education class is a priority in the registration process.

One respondent described a successful change to her college's intake assessment process. A full description follows, in the hope that others can benefit from this idea.

In the past, many low-level students would be put into the group testing situations using the *CAT* tests. They would find this experience very humiliating because they did not have the necessary skills to cope with

the testing material. Now, any student who does not have the prereqs to get into a course or program must sign up for a PrEP session (Preparation for Educational Placement). Through this 15- minute interview, we are able to determine which is the best way for the student to be assessed (i.e., one-on-one using CARA; group using CAT; ASD (access for students with disabilities); or ESL for those whose English skills are fairly low). We have had great success with the PrEP sessions, and we now have over 80 percent of students going through assessment registering in courses. I cannot think of anything else that I would wish for.

This college has implemented a procedure that reduces students' risk of humiliation during the assessment situation. While this procedure does not postpone the assessment process, it does create a more positive experience for the student and helps to alleviate test anxiety.

The data clearly indicated that intake assessments were administered more frequently than were on-going and exit assessments. Among the 400 respondents, 91 percent conducted intake assessments, 71 percent on-going, and 47 percent exit. The instructors wanted the opportunity to measure progress, particularly through on-going⁴ assessments, “on an as-need-basis, instead of an as-time-allows basis.” In order to measure progress in a reliable manner, the respondents noted that assessment tools need to have parallel forms for pre- and post- testing. A few respondents noted that a tracking or record-keeping system would assist in documenting and monitoring progress.

The percentage of respondents that conduct exit assessments is quite low, given that funders' accountability frameworks usually require programs to demonstrate learner progress. Yet, when we consider that many students in community-based programs “just drift away” and those in post-secondary programs leave with little or no notice, the statistics make sense. Studies have shown that students often drop out of programs because of socioeconomic-circumstantial factors (Long & Middleton, 2001; Roussy & Hart, 2002).

⁴The terms “formative” and “continuous” were also used to describe on-going assessments.

The Assessment Product

The respondents spoke of the qualities they wanted in an assessment tool. Data analysis revealed four commonly cited qualities: useful, user-friendly, current, and culturally sensitive.

Useful.

Many respondents were searching for the “perfect” assessment tool—a “foolproof instrument with 99.9 percent accuracy in results.” According to one respondent, this tool will “guarantee that my initial placement and individualized instruction will always be right for the student in question. Regardless of what assessment tool I use, there is always an element of hit and miss.” The findings indicate that respondents want reliable and diagnostic intake tools that determine placement and inform instruction, thereby optimizing teaching and learning. Instructors want on-going assessment tools that reveal how the students are doing and what to do next. They want assessment instruments to yield useful data that will “mean something” to instructors, students, and funders.

User-friendly.

The respondents emphasized that they wanted a user-friendly assessment tool—one that was simple to administer, score, and interpret. The need for a simple, easy-to-use tool appears to stem from two primary factors: time and expertise. For example, many of the instructors in post-secondary institutions assess students during class time, making the need for a user-friendly tool a necessity. While 80 percent of the survey respondents held a bachelor’s degree or higher, only 56 percent had taken a credit course focusing on assessment.

Current.

A common request on the respondents’ wish list was for updated assessment tools relevant to the curriculum and the student population. The findings show that the most frequently used standardized assessment tool—the *Canadian Adult Achievement Test (CAAT)*—was published in 1986 and has not been revised. One respondent, who coordinates adult basic

education programs for a school district that uses *CAAT*, expressed these concerns with older tests:

1. Sometimes they no longer match a curriculum that is relevant to the students' needs.
2. Sometimes the teacher modifies the curriculum to match the test.
3. Students may have access to old copies of the tests (or to students who have taken it previously), bringing validity into question.

In addition to these three points, older tests are usually based on outdated reading theories. *CAAT*, for example, is based on the text-based model of reading, rather than on a social constructivist or new literacies model. In fact, in spite of changes in reading theories, there has been little change in either the basic content or the format of standardized assessments since the 1930s (Bainbridge & Malicky, 2004).

Culturally sensitive.

Due to the diversity of students attending adult basic education programs, instructors want to use assessment tools that are “fair” and without “bias.” The respondents stressed that all tools need to be geographically and culturally sensitive, with respect to First Nations populations and visible minority groups who have taken English as a Second Language. Many students reside in remote areas, which means that they experience test items that are geographically biased. For example, consider a *CABS* test item that asks questions about paying parking tickets. Would this be relevant to students who live in isolated hamlets in the territories or in rural areas where parking tickets are non-existent? However, according to Johnston (1998), bias is always embedded in assessments. Johnston writes that “because of the cultural nature of literacy, it is not possible to create an unbiased literacy test; tests always privilege particular forms of language and experience” (p. 98). Despite Johnston’s claim, test developers are not off the hook when it comes to developing culturally sensitive assessment tools. Test developers have a responsibility to reduce bias in tests by analyzing item data separately for different populations and then identifying and discarding items that appear to be biased.

Human Resources

Capacity.

According to Merrifield (1998), in order to meet the demands of accountability, delivery agencies that provide educational services need the capacity to perform—that is, to achieve the performance goals, and to be accountable—that is, to document achievements. The findings from this survey indicated that many respondents are mandated by funding agencies to conduct comprehensive, on-going, and exit assessments, yet they do not have the capacity to fulfill this mandate.

While many respondents wanted more release or paid time for existing staff to administer and interpret assessments, others wanted to hire one person to conduct initial, on-going, and exit assessments. In fact, respondents from seven jurisdictions expressed a desire to have one person responsible for assessment. These quotations reflect the need for a new, expanded organizational capacity in community-based programs, colleges, and school boards:

Assessment would be done by one person whose job it was to assess the student, develop a learning plan with the student and evaluate progress on a regular basis – at least every three months. That person would have regular session with the tutors, to set learning goals and make sure the students were supported in the endeavour to achieve these goals. (Community-based program director)

I would have the money to pay one person to be our assessment expert. She would have time to spend outside the classroom with each learner for initial assessment and training plan development. (ABE school district director)

There would be an individual who devoted time and energy to assessing, evaluating, counseling, and supporting students in their development—fulltime—only job—work with, guide, support students. (ABE college director)

Fifty percent of the respondents who wanted one person to conduct assessments were from Ontario. Every month, literacy and basic skills (LBS) agencies in Ontario

are required to gather and input statistical information into the ministry's electronic information management system. While this system assists the ministry in determining whether programs are achieving goals and objectives, it presents a challenge to the LBS programs, as practitioners spend a great deal of time gathering and inputting information. In fact, the issue of capacity was cited as a recurring barrier with respect to implementing performance measurement in a survey that was conducted by the Ontario Literacy Coalition (2002). This might explain why so many Ontario respondents expressed the desire to assign one person to the role of assessment.

Communication channels.

Some colleges and school districts do have a testing or counselling centre where one person is assigned to administer intake assessments. In a few instances, colleges have a person within the adult basic education department (ABE) who is responsible for assessment. A few of the respondents who worked in testing or counselling centres expressed the need for more consultation with ABE instructors in order to ensure individualized instruction, based on the assessment. On the other hand, some of the instructors wanted the assessors in testing or counselling centres to arrange case conferences and to share test results with faculty in the form of teaching and learning strategy recommendations. One respondent stated that "there needs to be more discussion about potential students between the assessment officer and the instructor and/or chair who does the interviewing." This suggests that having one person assigned to assessment doesn't always ensure that instructors and students will receive the information they need to teach and learn. In addition to having an assessment or counselling centre, post-secondary institutions also need effective communication channels between assessors and instructors.

Referrals.

Usually, adult basic education practitioners do not have the qualifications to diagnose students at risk of learning disabilities. Consequently, many respondents want the financial resources to access experts to conduct psycho-educational assessments, or they want sufficient resources to contract professionals to determine specific learning

requirements and challenges. In summary, programs need the resources to make referrals when specialized assessments are required.

Material Resources

The majority of respondents spoke of access to assessment tools and professional development in the same breath. Choosing appropriate assessment tools can be a daunting task. First, one needs to know what is available in terms of commercial assessment tools. Second, funds to purchase these tools are required. Following the purchase of new tools, educators must deal with the next hurdle—learning to use the instrument. The complexity of the assessment instrument will dictate the amount of training educators will require in order to ensure accuracy and reliability during administration, scoring, and interpretation. According to the survey findings, respondents need the material resources of time and funding to access assessment tools and professional development.

The respondents expressed a desire for a resource library of assessment tools or access to a diverse range of materials. One respondent from a community-based program lamented, “I realize all the resources that are available but the time to study and implement them just is not available given the hours the program works on and the other needs that must be slotted into those hours.” Practitioners need time to explore and familiarize themselves with other resources.

The respondents want training to gain or enhance their knowledge about specific assessment tools; to learn about recent studies on assessment theories and methodology; to receive confirmation that their assessment practices are adequate; and to ensure that they “haven’t developed any bad habits or short cuts.” They also expressed a desire for networking or sharing sessions with their colleagues to “discuss and share resources pertaining to assessment.” Specifically, the respondents want to learn about the range of assessment tools that are “on the market, what they use, how they use them and when, what are the best tools to use to determine reading levels, writing levels, and math levels.” Finally, some respondents wanted mentoring to assist them with interpreting test results.



